

INDICATORS OF HOUSEHOLD INCOME FOR USE IN THE EVALUATION
OF AGRICULTURAL DEVELOPMENT PROJECTS

A.I.D. EVALUATION OCCASIONAL PAPER NO. 33
(Document Order No. PN-ABC-902)

Beatrice Lorge Rogers
Tufts University

U.S. Agency for International Development

October 1989

The views and interpretations expressed in this report are those of the author and should not be attributed to the Agency for International Development.

TABLE OF CONTENTS

Foreword

1. Introduction: Why Measure Income?

2. Income Indicators, Direct and Indirect

2.1 Wealth

2.2 Income

2.2.1 Cash

2.2.2 In-Kind Payments

2.2.3 Time of Household Members

2.3 Expenditure

2.4 Consumption

2.5 Quality-of-Life Indicators

2.6 Choice of Indicators

3. Sources of Data

3.1 Secondary Data

3.2 Administrative Records

3.3 Project Records

3.4 Primary Data for Evaluation Purposes

4. On-Site Data Collection Methods

- 4.1 Avoiding Bias in Data Collection
- 4.2 Approaches to Data Collection
 - 4.2.1 Direct Observation
 - 4.2.2 Focus Groups
 - 4.2.3 Key Informant Interviews
 - 4.2.4 Small-Scale Surveys

Appendix. Sample Survey Questions

References

FOREWORD

The Agency for International Development (A.I.D.), Bureau for Program and Policy Coordination/Center for Development Information and Evaluation (PPC/CDIE), in cooperation with the Bureau for Science and Technology and three regional bureaus, organized a workshop on indicators for measuring changes in income, food consumption and food availability, and the natural resource base. The purpose of the workshop was to identify and discuss a set of simple, practical indicators that can be used by overseas Missions and A.I.D./Washington for monitoring the impact of agricultural and rural development assistance.

The workshop was held on June 20-22, 1988, in Virginia and was attended by 60 development specialists, including A.I.D. staff, consultants, and outside experts. Four background papers written by experts were presented at the workshop; this paper is one of them. The titles of the others are "Impact Indicators for Measuring Change in Natural Resource Base," "Food Availability and Consumption Indicators, and "Impact Indicators: General Issues and Concerns."

This report by Krishna Kumar, entitled "Indicators for Measuring Change in Income, Food Availability and Consumption, and the Natural Resource Base," presents the main findings and conclusions of the workshop. In addition, CDIE has issued another paper, related to the workshop topic, entitled Methodologies for Assessing the Impact of Agricultural Development Projects, A.I.D. Program Design and Evaluation Methodology Report No. 11.

I am confident that these publications will be of great help, not only to A.I.D. staff and contractors but also to host governments and institutions struggling to develop effective and efficient monitoring and evaluation systems for development activities.

Janet C. Ballantyne
Associate Assistant Administrator
Center for Development Information
and Evaluation
Agency for International Development

October 1989

1. INTRODUCTION: WHY MEASURE INCOME?

Indicators of household income are used for several purposes in the development and implementation of agricultural interventions. First, income levels indicate where, and specifically among which households, an economic or agricultural intervention is needed. That is, information on income can be used as a basis for targeting communities and households in project design. Income information is also important for identifying resources available within target communities and households -- information that is critical for assessing the feasibility of project plans. Finally, information on changes in household income is central to the evaluation of program and project success. The intent of agricultural development projects, as of any economic development intervention, is to improve the standard of living of poor households and household members. Indicators that show changes in income let planners and managers of agricultural development programs know whether the programs are having the desired effect. And when data collection on income indicators is carefully undertaken, the information can be used as a guide to modify program design when program outcomes do not appear to be entirely positive.

The focus of this discussion is on income indicators that can be used to evaluate the effectiveness of agricultural interventions and, when necessary, to provide feedback for modifying the interventions appropriately. This focus on evaluation dictates that indicators measure income changes at the household level, and even at the level of the individual within the household. Improvements in community wealth and income can be misleading if the most economically vulnerable households have not benefited from the change, or have suffered a loss of income at a time when the majority or perhaps the more visible households have experienced gains.

Income information is useful for comparing the relative success of programs in different settings. Indicators of household economic status that are valid across countries and cultures are much more difficult to find than are indicators for comparing villages within a single geographic area or for comparing project villages before and after implementation of an intervention.

2. INCOME INDICATORS, DIRECT AND INDIRECT

Income is used as an indicator of the economic status of a household, and economic status is generally taken as an indicator of household welfare. Household income does in fact measure a household's access to resources, to the extent that these are purchased. However, income is limited as a proxy for welfare because of the following considerations:

- Many rural households have resources, or have access to resources, that have not been purchased for example, home-produced goods, unpaid family labor, and freely available goods such as wild food or building materials. To the extent that such resources are used, income alone will understate the economic status and the welfare of households.
- Uses of income are in part determined by who earns income and who controls it. Income increases the range of consumption opportunities available to a household, but whether or not this consumption will lead to an improvement in any particular indicator of welfare (such as health status, educational achievement, land ownership) cannot be determined by knowing only the level of income.
- Similarly, the uses of income are affected by the period in which income is earned, its reliability, and its form (cash, goods, services). For measuring effects on welfare, changes in the characteristics of the income flow may be important, as well as changes in the amount of income.
- Income is a flow; that is, income must be measured in terms of some fixed reference period. The level of income for any household, but perhaps especially for marginal, rural households, varies considerably by season and from year to year. This means that income measured at any particular point may not be an accurate representation of the household's welfare, unless saving and dissaving (i.e., buying on credit or drawing on savings) are taken into account.
- Income levels are generally not comparable across countries. Even if income levels are converted to a single currency equivalent using an appropriate conversion rate, the significance of the same income as an indicator of well-being will be different in different settings.

A variety of indicators should be examined in any project impact evaluation. In an evaluation study, several categories of indicators of well-being can be considered for measurement, including wealth (household assets), income, expenditures, consumption, and, for want of a better term, quality of life.

2.1 Wealth

Ownership of a set of assets can be a measure of household wealth. These assets can include productive

resources, such as land, farm implements, and irrigation systems, and consumption goods, such as housing, water supply, plumbing, household appliances, cars, and bicycles. (Some of these consumption goods can also be considered productive resources, at least in that they represent investments in human capital. For example, availability of piped water is associated with improved health status.)

There are several advantages to using assets or wealth indicators to assess economic status. Assets in general are readily observable. An interviewer or even an extension agent can readily observe whether a house is constructed of mud, brick or cement, whether the roof is metal or thatch, whether there is a refrigerator in the house. Questions about ownership are simple and concrete, even when assets are not readily observable (e.g., radios). Many difficulties in the measurement of household economic status arise not from respondents' unwillingness to answer questions, but from their inability to do so, either because they do not know the answer or because they do not fully understand the question. Unless a respondent has a motive for lying, he or she should be able to understand and answer questions about ownership of assets.

Another advantage of using assets as an indicator of economic status is that ownership of assets is much more stable over time than is income. Thus, changes in the ownership of assets reflect relatively longer term shifts in economic status. However, they may not be good indicators of short-term economic fluctuations, which can have important implications for welfare (e.g., seasonal shortfalls in income resulting in periodic nutritional deficiencies).

The measurement of asset ownership poses some difficulties, however. For example, the concept of ownership may not always be perfectly clear, especially with respect to land. There are assigned differently from rights to sell or pass on a piece of land; formal ownership (with written title) may be defined differently from traditional access, even though for most purposes both forms of ownership are equivalent. These culturally specific definitions of ownership must be fully understood, and a decision made about the appropriate one to use, before ownership can be used to measure welfare for project evaluation purposes.

Land ownership as an indicator of economic status poses other difficulties as well. Landowners may be reluctant to reveal the amount of land they own because they fear taxation or confiscation of their land because of land reform. Also, land, while very important, is not the only source of wealth, even in agricultural communities. Wealthy merchants and moneylenders, for example, might not own land. Thus, land can never be taken alone as an indicator of economic status. Other possessions, such as quality housing or transportation, may be more universally applicable indicators of economic status.

Another concern about using asset ownership as an indicator of economic status is that indicators of wealth are highly culture specific. It is absolutely essential not to make any untested assumptions about what assets are valued in a given location. For example, a survey in rural Pakistan found that indoor latrines, far from being an indicator of wealth, were owned only by households that were so poor that they were clustered together. In wealthier households, latrines were not used. In Mali, access to piped water was not a reliable indicator of wealth: in some areas, a well in the compound was preferred to piped water just outside the house (personal observation).

While there is no substitute for a thorough knowledge of the communities under study, and no data-collection effort should ever be undertaken without adequate pretesting, there are statistical techniques, such as principal components and factor analysis, that permit testing after the fact to see whether all the presumed indicators of wealth indeed group together.

Despite the concerns discussed above, asset ownership information is one of the least expensive types of data to collect. In many cases, data on asset ownership can be collected by simple observation with minimal intrusion into households. And data can be aggregated to the community level so that, for example, the number of households owning houses with tin roofs in a base year can be compared with numbers for a subsequent year of the project, or similar data on a project community can be compared with data on a suitable control.

2.2 Income

Income represents a flow of resources to a household or an individual over a specified period of time. Income can be defined in terms of cash received, but such a definition is likely to leave out important resource flows, especially in rural areas.

Real income is defined as the flow of cash and nonpurchased goods and services, measured in terms of their monetary value. Recently, the concept of full income has been introduced in the literature on measurement of economic status. Full income includes cash received plus the monetary value of goods and services and the time of household members. The inclusion of household members' time in a definition of income is important, because members' time represents a valuable contribution to the welfare of households. Cash and physical resources can be conserved to the extent that household labor is substituted for purchased inputs in household maintenance.

In some cultures, income is a highly sensitive subject and questions about income are considered intrusive. In such cases, it is necessary to find indirect indicators of economic

status, such as asset ownership, expenditure, consumption, and direct measures of welfare. In any case, as mentioned earlier, income itself may not be the most valid indicator of true economic status (i.e., access to consumption goods), which is what an evaluation of agricultural interventions seeks to measure.

The biggest advantage of obtaining information on income per se is that the information permits an analysis of the relative importance of different income sources (e.g., agricultural versus nonagricultural, females versus males or adults versus children, earned versus unearned) and of the reliability and period of income. Such information is important. An intervention that does not raise agricultural income but that, through new technology perhaps, frees household members to take wage employment, may significantly improve household welfare. Or an intervention that evens out the flow of income during the year, perhaps by introducing a new crop variety that has a different growing season, may raise a household's economic level because of increased income stability even if annual income is unchanged. Finally, an intervention that alters the flow of income among members may significantly alter the consumption level of the household, because different members' preferences may be served.

No matter how carefully a survey has been designed to measure income, it is unlikely to achieve exact accuracy. Errors are due to lapses in memory, deliberate misstatement, and fluctuations in income that do not reflect real changes in economic status. For the purpose of assessing only the relative importance of different income sources, such error may not matter much, as long as it does not involve bias -- that is, as long as all sources of income are equally subject to errors of reporting.

Measurement of cash income is complicated by the choice of a reference period, the need to ask separate questions about each possible source of income, and the need to question earners individually about their earnings. Perhaps the most intractable problem in measuring cash income is that, in places where income is received irregularly, in small or in very variable amounts, and without written records, people simply do not know, or have difficulty remembering, the amount earned in the past month or year.

In most studies in agricultural areas, the appropriate reference period for measuring income is a year, which covers the agricultural cycle. Different income sources dictate different reference periods, though. Wages should be measured in terms of the pay period (month, week, day), with appropriate questions asked about the number of periods worked during the previous year. Other sources, such as informal transfers from family members, and formal transfers, such as social security, should be measured with respect to the period in which they are normally received. Moreover, yearly variation in income can be substantial, so that any given year is only an approximate indicator of permanent economic status. For example, if cash income from

agricultural sales (or from agricultural labor) is low because of poor rainfall in a given year, the lower income should not necessarily be taken to indicate failure of the agricultural intervention being evaluated.

In an investigation measuring income, an interviewer needs to identify explicitly each possible income source, because respondents easily may forget income sources that are variable or very irregular, or sources from which they have received no income recently. Each earner has to be questioned individually about his or her earnings, preferably in private, because it is common for household members to keep their earnings secret from each other. It is also important to probe respondents on earnings, especially women. It is now well recognized that women often report that they do not work for pay, because they primarily identify themselves as homemakers; yet they may have significant earnings from casual businesses (selling homemade goods or produce, for example), or even from wages. Children can also be important providers of income in low-income households.

The interpretation of income information is problematic for a variety of reasons. As was mentioned above, not all income is used in the same way for household welfare. The income of older children, for example, may not be contributed to the general resources of the household to the same degree as that of the adults in the household. It has been suggested that women's income is used differently from men's income: specifically, that men are more likely to use income for long-term investment and women are more likely to spend income on basic household consumption needs (especially food). Although this is a controversial point, the importance of knowing not only the amount of income but who earns it is clear. Also, income received in small, regular increments is more likely to be spent on short-term consumption, while large, irregular payments -- e.g., income from crop sales -- are more likely to be spent on large, indivisible expenditure items such as school fees and investment. In sum, many aspects of income, in addition to its level, need to be considered when interpreting income as a direct indicator of welfare.

2.2.1 In-Kind Payments

The problem of measuring in-kind income is that of evaluating it in monetary terms. This problem is perhaps most easily resolved in the case of food, for which the purchase price (for households that do not produce food either for sale or for their own consumption) or the farmgate price (income forgone for producers) can be used. Evaluating in monetary terms the value of payments received in the form of prepared meals, clothes, housing, and so on is more difficult, partly because variations in quality are difficult to assess, and the recipient's loss of choice regarding the use of the income should somehow be taken into account.

The income period and earner are as important for measuring in-kind income as they are for measuring cash income, and the problems of choosing a reference period are the same. Monetizing the value of home-produced food poses another problem. Many households derive a substantial percentage of their food from home production, but they cannot estimate an annual monetary value for the quantity of food consumed because they collect from their fields or gardens only the quantity they need for the day's meals. For example, the only way to estimate the income value of such items as fruits, vegetables, tubers, and eggs and milk, which do not have one short, specific harvest period, is to measure actual consumption in a given reference period and estimate its annual value, taking seasonal variation into account if possible. (This can be done either by measuring consumption at several points during the year or by asking questions about the seasonal availability of the most important foods.)

2.2.2 Time of Household Members

Monetizing the value of household members' time poses an intractable theoretical problem of measurement. One approach measuring the monetary value of time has been to estimate the market wage rate commanded by people with characteristics (age, gender, educational level) similar to those of the member in question and use that wage rate as an estimate of the income forgone as a result of that person engaging in unpaid home production. However, this approach does not account for the account for the self-selection of nonworking household members, nor for the fact that wage labor involves fixed inputs of time in a fixed location, so that its relevance to relatively unstructured time at home is questionable. Also, clearly not all time in a 24-hour period can be valued the same way.

There is no simple solution to the problem of estimating the monetary value of home labor. For the purposes of an evaluation, it is important to record the number of working and nonworking (for pay) adults in a household and to recognize, in interpreting the data, that households of equivalent size and real income are better off, in terms of resources available to promote family welfare, if they have more of members' time available. That is, the availability of time does not necessarily have to be measured in the same terms (monetary units) cash and in-kind income, but it must be taken into account in assessing a household's standard of living.

2.3 Expenditure

Household expenditures are often easier to measure than is household income for several reasons. First, respondents

are likely to know and to remember how much they spent on something. Second, expenditure may be a less sensitive subject than income. In many places, direct questions about income are considered intrusive, but people willingly report on their expenditures. Third, the permanent income hypothesis holds that expenditures are more stable than is income: as income fluctuates, households will save or dissave in order to maintain a level of expenditure that matches their longer term expectations of their economic status. However, it may be argued that in very poor households, daily expenditures are determined by short-term income. Poor households, so the argument goes, do not have the luxury to save or to dissave, so that their expenditures are not more stable than their income. In fact, this argument is not borne out by the evidence of widespread use of credit in poor communities. Nevertheless, even if expenditures were not stable, it would mean that expenditure is no better than income as an indicator of economic status, not that it is worse. One can also argue that expenditure is more closely related to the underlying concept of economic level than income, since expenditure represents goods actually consumed by the household, whereas income represents potential consumption.

The problems of choosing a reference period in the measurement of expenditure are similar to those encountered in income measurement. Information on expenditures for an item should be based on a logical reference period. For example, transportation costs may be weekly; rent, if it is paid, may be monthly. Food may be a daily or a weekly cost, depending on local or household patterns. In some cases, the reference period for food staples may be as much as a year if people are in the habit of buying one large sack of grain for longterm use. This is probably less characteristic of very poor families, who are unlikely to have the cash for such large purchases. The interpretation of purchases of consumer durables as indicators of income also poses problems. If large numbers of households in a community are beginning to invest in such items as cars or major appliances, it may be an indication that the general level of wealth in the community is rising, which may, in turn, be an indicator of program impact. But the purchase of such items cannot be added to regular periodic expenditures and included in an estimation of expenditure as a proxy for annual income because such investments are not likely to be repeated and may represent several years of consumption forgone (either in the past, if the item was paid for out of savings, or in the future, if the item was purchased on credit). Such purchases should probably be analyzed separately from more regular expenditures.

2.4 Consumption

Expenditure measures only items purchased for cash or on credit and thus, as when cash income is measured, may leave out major categories of consumption obtained from home production or from gifts and in-kind payments. A study of household economic status, particularly if conducted in rural

areas, must include nonpurchased consumption goods.

Measuring the value of nonpurchased consumption goods requires assessing the market cost of the items as though they had been purchased. In the case of food consumption, for households that produce the food they consume, the food should be valued at the farmgate price. Such an estimate would best represent the level of consumption of other items they could have purchased if they had sold the food. For households that obtain food as payment or gifts, the food should also be valued at the retail price. In many rural areas, many foods are sold directly by producers to consumers, so the farmgate and the consumer price are the same. In the case of housing, many rural communities have no rented housing, and ownership simply means finding free materials and building on unoccupied land. In such cases, it may not be meaningful to monetize the value of nonpurchased housing.

Reference periods for measuring consumption, like those for expenditure, should be chosen to suit the item being measured. For food, an appropriate reference period may be a few days or a week, but seasonality also needs to be taken into account. Food is both cheaper and more available for home consumption during some seasons than others. Seasonal festivals involving feasts or fasting (such as Ramadan, Eidh, Christmas) should not be included in a reference period intended to represent the household's normal level of food consumption.

In general, food consumption is a special category of consumption, posing specific measurement problems. In rural areas where subsistence farming is a significant activity, food purchases cannot serve as a proxy for food consumption. Instead, actual household consumption must be estimated, which is probably better done through recall than through direct weighing or measuring of food. Recall using food models to assist in estimating quantities is less accurate, perhaps, than weighing. But the intrusiveness of food weighing is certain to cause significant distortions in behavior, making the "accurate" measurements invalid as indicators of usual consumption.

Reference period is also important when measuring food consumption. One day's consumption is not an adequate indicator of a household's usual consumption level. At least 3 days are necessary to obtain a reliable estimate of caloric and protein intake. (More days are necessary to estimate consumption of micronutrients, but for the present purpose, estimation of calories and protein is sufficient.) Seasonal variation may also be significant and should be taken into account in the design of an evaluation.

The particular advantage of food consumption as an indicator of economic status is that it is relatively culture-free. Caloric and protein adequacy have the same meanings in all settings. Furthermore, there appear to be

some universal trends in food consumption patterns associated with economic status. For example, as incomes rise, first the level of caloric intake increases. With further improvements in income, the proportion of calories and protein derived from animal sources tends to rise. The variety of items in the diet also appears to be positively related to income level across cultural settings.

Since food is one consumption good that is equally important in all populations (in contrast to housing, modes of transportation, or durable goods), the measurement of food consumption can provide an indicator of economic status that, while approximate, is likely to be cross-culturally valid.

2.5 Quality-of-Life Indicators

Perhaps the most indirect measures of income, but also the most direct measures of welfare, are indicators of the physical quality of life, such as nutritional status, health status, morbidity, and mortality. These indicators are closely related to income and expenditure levels, but they are determined by many factors other than economic ones. Nonetheless, quality-of-life indicators should be included, when possible, in evaluations of project impact, because they provide information directly related to the welfare goals that represent the underlying purpose of economic development initiatives.

The nutritional status of children in households is relatively easy to measure. There are widely recognized height-for-age, weight-for-age, and weight-for-height standards that can be used for assessment. Generally, height for age is interpreted as an indicator of long-term dietary adequacy, and weight for height is interpreted as an indicator of changes in the short term. The nutritional status of children is one indicator of the availability of resources at the household level, and changes in the prevalence of malnutrition are probably good indicators of changes in the level of welfare of a community over time.

Infant and child mortality are also indicators of resource availability and welfare at the community level, although they are not suitable for assessing household-level changes (because the frequency of such mortality is low).

Quality of life may also be indicated by variables such as the number of children in school at a given age and the amount of leisure time available to adults. These are both indirect measures of the severity of need for household labor. As a household's economic level improves, it is better able to spare labor from income-earning activities and spend time on consumption (leisure) and investment in human capital (education).

2.6 Choice of Indicators

Table 1 summarizes the relative advantages and limitations of various indicators of household economic status discussed in this paper. Each indicator measures something specific, and each can be considered as part of an impact evaluation. A number of measures discussed can be interpreted at the community and the household level, and both levels should be included in an evaluation. Community-level changes are a powerful indicator of the success (or failure) of a project. There are cases of economic change in which improvements in communities' standards of living were obvious to casual observers simply from the number of tin roofs, electrical connections, and television antennae on houses. However, there are also cases in which the visible change in the general level of economic status masked the marginalization of some categories of households that did not benefit from the changes. One study even found that the

Table 1. Advantages and Disadvantages of Selected Income Indicators

Indicator	Advantages	Disadvantages
1. Wealth and Asset Ownership	Concrete and easily measured Often directly observable Relatively stable over time Aggregation to community level is possible if sampling procedures are followed	Concept of "ownership" is culture-specific; must distinguish use rights, control over allocation, etc. Fails to reflect short-term fluctuations in economic status Deliberate understatement of assets due to fear of taxation is possible (especially for land) Assets that are valid indicators of economic status are culture-specific
2. Income	Permits measurement of the relative contributions of	In some cultures, questions about

different income sources and earners income are unacceptable

Permits measurement of the frequency, timing, and reliability of income streams Income may not accurately represent consumption level if savings and dissavings are not accounted for

Permits measurement of returns in labor (wage rates), important for assessing value of time Choice of reference period greatly affects measured levels of income. Income fluctuates both within the year and between years

The larger the proportion of in-kind income received, the more difficult it is to quantify in monetary terms

Evaluation of home-produced, home-consumed food and of the time of household members poses special problems

Individual interviews with all household earners required

3. Expenditures In many cultures, expenditure is less sensitive than income Choice of reference period affects measured level of expenditure, though possibly Expenditures are easier to remember than irregular income less than for income

Expenditure on consumer durables may give misleading estimates of

Expenditures fluctuate less over time than income and are a more stable indicator of economic status. To the extent consumption is not purchased, expenditures understate economic status.

Expenditures represent actual not potential consumption and thus reflect welfare better than income.

4. Consumption More accurate reflection of welfare than cash expenditure Choice of reference period affects level of measure

Has all the advantages of cash expenditure. It is often difficult to impute a monetary value to non-purchased goods.

Home-produced, home-consumed food is difficult to measure.

5. Quality-of-Life Indicators Concrete and easy to measure based on anthropometry than household economic status May be affected by factors other than household economic status

- a. Nutritional Status Concrete and easy to measure based on anthropometry (for children) adults Not easy to measure for

Direct indicator of welfare

Can be aggregated to community level if sampling procedures are followed

- b. Infant and Child Mortality Direct indicator of welfare household income is very indirect Relation to

Suitable to comparison of communities, but not individual households

apparent change in household economic status believed to have resulted from an economic development project had been due to changes in the types of households living in the community: those who had benefited from the project had stayed, but many households had simply moved away and had not been part of the follow-up study.

Consideration of indicators relating to individual household members is an important element in a project evaluation. Changes in the flow of income among earners is not only an issue of how income is used (as discussed above), it may also be an issue of equity if some household members lose income or access to resources while others gain.

3. SOURCES OF DATA

The scale of the intervention and the scale of the evaluation study will certainly determine the range of possible data collection methods and sources of data that can be used in an assessment of program impact. Possible sources of data are as follows:

- Secondary data, including population censuses, agricultural censuses, household income and expenditure surveys, and health and nutrition surveys
- Administrative records of government and nongovernmental agencies and programs
- Data obtained as part of the administration of the agricultural intervention project being evaluated
- Observational data collected as part of an evaluation study
- Interview data obtained from a household survey that is part of the evaluation study

3.1 Secondary Data

Many countries have data available from large-scale surveys, including censuses, income and expenditure surveys (typically conducted to develop a basis for a consumer price index), and special-purpose surveys. For several reasons, the value of such data sources for evaluation of a specific project is likely to be limited. Access to the data may be difficult to obtain, because the institution responsible for the study may not wish to release it. Also, the level of detail in such studies is rarely sufficient to permit the type of analysis needed for project evaluation purposes. Furthermore, even though the costs of data collection have already been incurred, the additional costs, in terms of time and money,

required for cleaning and analyzing large survey data sets are not trivial. In most countries, these large-scale surveys are usually not performed frequently enough to be useful for evaluation purposes. And, finally, the quality of the data collected must be a concern for any user not directly involved in the design and supervision of the data collection effort.

Nevertheless, investigating the availability and usefulness of secondary data sources in a given setting is still worthwhile, because there may be cases in which none of these problems exists. Surveys can be valuable if the data are good and if the level of detail is sufficient to enable disaggregation of the data by appropriate region. Such data sources are most likely to be useful as baseline data with which subsequent changes brought about by a project can be compared. Information on expenditure levels, consumption of different classes of goods (as an indicator of economic status), type of dwelling, and ownership of assets is likely to be useful to an evaluation.

3.2 Administrative Records

The main concerns with using administrative records as sources of data for impact evaluations are those of bias in the population reflected in the records and systematic errors of measurement due to incentives for misreporting. For example, records of landholdings may systematically exclude owners of the smallest holdings, or those whose title or access to the land is informal or questionable. The size of landholdings may be systematically understated if the information is used for tax purposes. Data on the use of agricultural inputs may include only households regularly visited by the agricultural extension agents and thus be inaccurate; these agents may have an incentive to overstate the use of modern inputs if their own effectiveness is being judged by the number of adopters of modern methods. Production and yield information may also be inaccurate if producers withhold some of their product for sale through unofficial channels.

Once again, administrative records may provide useful information on changes in an area, but the user of such information needs to be aware of the possible pitfalls.

3.3 Project Records

It would be ideal if agricultural projects could have a built-in process for collecting information useful in assessing the income changes brought about by the project. However, care must be taken not to burden project administrators with tasks not directly relevant to what they perceive as their jobs. Indicators directly related to the project's own activities, such as irrigation wells dug, pumps installed, or tractors purchased, may be useful indicators of changes in the economic level of households. But it is unlikely that agricultural project workers will be inclined to keep track of household food consumption, children's nutritional status, or the time use of individual members of agricultural households.

Nor are project officials likely to be the most suitable people to collect that information.

Those responsible for designing agricultural projects may feel that building in a separate ongoing evaluation component would represent a diversion of project resources from the primary goal of project implementation. However, such evaluation would not necessarily have to be separate from the implementation activities. Tracking the effects of an agricultural project on the income, consumption levels, and welfare of households and their members is highly relevant to implementation. Evaluations can identify problem areas early so that any negative effects can be forestalled.

However, monitoring the project's effects on household income in the project sites poses a serious conceptual problem because such an effort does not include an appropriate comparison group. As an evaluation design, monitoring provides the possibility for before-and-after comparison, but it does not provide for a control group as would be done in a formal evaluation.

3.4 Primary Data for Evaluation Purposes

Secondary data may provide some useful indicators of agricultural project impact if evaluators are lucky and suitable data are readily available. However, in most cases a reliable evaluation requires a focused effort, including the collection of primary data. This does not necessarily imply a massive data-collection effort. A judicious combination of direct observation, informal interviews, and possibly small-scale surveys can obtain the necessary information efficiently when the effort is carefully planned and where it makes use of existing knowledge and experience in the field.

It is worth pointing out that too much emphasis on avoiding the expense of primary data collection is false economy. It is rare to find sufficient sources of secondary data to assess the impact of a particular agricultural intervention in a given region. Moreover, relying on secondary data alone may give inadequate or, worse, misleading information on the effectiveness of the program. However, well-planned evaluation studies can provide the needed information and, even more important, a basis for understanding what aspects of the project, if any, need to be modified. A small-scale study performed relatively frequently is a more ideal structure for an evaluation study than a onetime, large-scale study. One reason for this preference is that a low-cost but ongoing evaluation study can provide data useful for modifying the program, which is clearly more useful than a large final evaluation that simply points out whether the project had the desired impacts -- when it is too late to do anything about it.

Several researchers have suggested approaches suitable for relatively rapid, low-cost data collection (e.g., Scrimshaw and Hurtado 1987, Kumar 1987a, Rogers 1988). One such approach, taken largely from Rogers 1988, is presented in the following section.

4. ON-SITE DATA COLLECTION METHODS

Direct observation and formal and informal interviews with local informants at the site(s) of the project and at suitable control sites are essential to a reliable evaluation. Direct observation should be used whenever possible to complement and validate the information obtained from group and individual interviews. Also, experienced social scientists and development professionals, including local professionals when feasible, should be involved in performing such evaluations.

4.1 Avoiding Bias in Data Collection

Data collection for project evaluation may be on a smaller scale than a survey, and the techniques may be less formal. However, the principles of sampling are essential to ensure that a representative range of respondents is studied. In survey research, bias in the selection of respondents can be avoided by applying proper sampling techniques to ensure that every member of the population being studied has a known, non-zero chance of being observed. The population being studied may be households, farmers or other members of the community, extension agents, or a variety of other units.

Observations should be conducted at randomly selected locations and times of year (if possible; if not, recall should be used to account for seasonal variation). Households or other units should be selected from the full range of geographic locations and be characteristic of those in the project area. Although it is too costly and time-consuming to draw up an exhaustive sampling frame, it is still possible to ensure adequate sampling of the population under study. For example, the target region can be divided into zones according to distance from the agricultural extension station or the cooperative to ensure sampling some households from each zone. When the relevant household characteristics are not geographic but relate to landlessness or male/female headship, for instance, care must be taken to include some representative households from each subgroup (an approach known as "quota sampling"). Sampling of observations based on convenience of time and place is bound to be nonrepresentative and to provide misleading results and thus should be avoided at all costs.

4.2 Approaches to Data Collection

Several well-accepted methods of data collection for project planning and evaluation are designed to be relatively low in cost and to provide results relatively quickly (cf., Kumar 1987a, 1987b; Scrimshaw and Hurtado 1987; Rogers 1988). Different methods are suitable for different types of information and are described briefly in the following paragraphs. Generally, such data collection requires trained and experienced people to work in the field, and the use of such experts makes the data collection more cost-effective.

4.2.1 Direct Observation

Direct observation of public behavior permits the observer to validate information reported in the literature or by local informants. A format for collecting the data should be developed specifying the number of observations required and the procedure for obtaining them without bias. When possible, observation should be in quantitative terms. Structured data collection instruments for direct observation should specify the precise information required. Examples include the number of women and men attending agricultural extension classes in a sample of extension sites; relative frequency of men and women, and of different age groups, performing a given (publicly observable) task, such as weeding or land clearing; time required to perform a given task; and what tasks are performed together or in a fixed sequence.

4.2.2 Focus Groups

A focus group is composed of a small number of people (about 6 to 15) with similar interests in a project; the group meets with a discussion leader who guides a 45-minute to 2-hour discussion on a particular topic. (See Kumar 1987a for a fuller discussion.) Examples of such groups might be small landholders growing a particular crop or members of a marketing cooperative. The idea of focus groups is to get discussion going among the participants rather than to conduct a question-and-answer session. The leader's job is to keep the discussion on relevant topics and to move the discussion along when a topic seems to have been exhausted. The leader should have a list of general questions that participants in the focus group can answer and in which they are interested. The questions should be designed to elicit information without suggesting responses. "What is involved in selling your coffee?" is a more appropriate opener than "Would you like the cooperative to provide transportation for your coffee crop?" Of course, focus groups must be conducted by a person who is fluent in the local language.

When several different groups have an interest in the project, then several focus groups are needed, because some people may be unwilling to discuss their opinions in front of people whose interests in the project diverge from their own. (For example, landless laborers may have a different interest in the project than small landholders.) Focus groups, by encouraging discussion, often reveal unanticipated aspects of a particular issue. Feelings, preferences, and attitudes may come out that could not have been anticipated but that might explain why a particular project outcome has been observed. Since discussions are public, people may correct each other's reports and improve the reliability of the information provided.

Of course, focus groups do not provide information about attitudes or behavior that people are unwilling to reveal to their neighbors. No form of data collection suitable to the project planning process will reliably obtain information on embarrassing or illegal activity, although a sensitive leader can often address relatively private subjects. Another

drawback of a focus group is that responses may reflect social norms rather than actual behavior, because people are responding in public. Furthermore, people may describe what they believe to be general practice, even when they know their own behavior in specific instances does not conform to it. This underscores the need for structured direct observation, when possible, to complement data obtained from informants.

Note taking in focus groups should be thorough but unobtrusive. When it seems inappropriate to take notes during the session, notes should be completed immediately after leaving the meeting. In a focus group, it can be helpful to have two people working, one guiding the conversation and the other taking notes.

4.2.3 Key Informant Interviews

Individual interviews using a relatively unstructured set of questions or topic guides (Scrimshaw and Hurtado 1987) can substitute for or augment focus groups if privacy is considered essential or if it is difficult to reach some people in a group setting. The questions are used to introduce a general subject area, and the respondent can answer, focusing on the aspects of the question most relevant to him or her. The advantage of such semistructured interviews is that, as in a focus group, the interviewer can pursue a line of questioning that leads in unanticipated directions. It is often advantageous to contact key informants several times over the period of the evaluation study. Ongoing contact reduces the reliance on any single interview, permits the development of personal rapport over time, and allows the interviewer to follow up on new information obtained from other sources.

4.2.4 Small-Scale Surveys

Surveys are distinguished from key informant interviews by the relatively larger number of respondents; greater use of closed-ended, precoded questions; and more rigorous application of sampling techniques. A survey must be the last stage in the collection process, because the key informant interview and focus group methods, as well as direct observation, are essential to ensure that the closed-ended survey questions are meaningful, cover all the relevant aspects of the problem, and are phrased in a culturally appropriate manner. The advantage of the survey approach is that it provides greater assurance of statistical representativeness. If the sampling method is indeed representative, then surveys permit evaluators to assess the frequency of particular situations or attitudes, not just to note their occurrence. Also, the more structured approach to interviews ensures that all questions are covered in all interviews in the same way. The possibility of bias due to the respondent's defining the focus of the interview is thereby avoided.

5. SUMMARY

Income is only one indicator of the economic status and welfare level of households and of individuals. A variety of other indicators is also suitable for measuring economic status, including asset ownership and wealth, expenditure, consumption, and direct welfare measures such as the nutritional status of children (a household-level measure) and child morbidity and mortality rates (a community-level measure).

Income is subject to problems of measurement, including the selection of an appropriate reference period, the fact that income is subject to severe fluctuations over the short run, and, in some settings, the reluctance of people to discuss their income. In rural and developing country settings, a definition of income must include income in kind as well as income in cash. Health (i.e., asset ownership) may be more stable than income as an indicator of changes in economic status. Expenditure and consumption (consumption includes cash expenditure and goods consumed from unpaid sources) are usually less sensitive topics than is income and may also be more stable and more direct indicators of welfare.

Measures of the effectiveness of programs must be sensitive to changes in the form, timing, and flow of income, as well as changes in the amount of income received by households.

A reliable evaluation is likely to require a primary data collection effort. Secondary data sources, such as existing surveys and administrative records, are unlikely to provide the level of detail necessary for evaluation purposes. Also, such data collection is unlikely to be performed frequently enough to be useful for project or program evaluation purposes. Secondary data, if they are available, easily accessible, and of satisfactory quality, may be useful as a baseline with which changes resulting from a project can be compared.

Frequent, small-scale project evaluations are preferable to larger, more expensive evaluations conducted less frequently, because ongoing evaluation permits the use of the data obtained for project modification. The cost and effort involved in project evaluation are likely to be more than repaid by improvements in cost-effectiveness and in the likelihood of successful outcomes of agricultural interventions.

REFERENCES

- Kumar, K. 1987a. Conducting Group Interviews in Developing Countries. Washington, D.C.: Program Design and Evaluation Methodology Report No. 8. Agency for International Development. April.
- Kumar, K. 1987b. Rapid Low Cost Data Collection Methods for A.I.D. Washington, D.C.: Program Design and Evaluation Methodology Report No. 10. Agency for International Development. December.
- Rogers, B. 1988. Incorporating the Intrahousehold Dimension Into Development Projects: A Guide. Paper prepared for Agency for International Development's Bureau for

Program and Policy Coordination, Office of Policy
Development and Program Review.

Scrimshaw, S., and E. Hurtado. 1987. Rapid Assessment
Procedures for Nutrition and Primary Health Care. Los
Angeles, California: University of California, Latin
American Center.